Cluster Analysis of 4-Day Back Trajectories Arriving in the Barcelona Area, Spain, from 1997 to 2002

Oriol Jorba and Carlos Pérez

Laboratory of Environmental Modeling, Escola Tècnica Superior d'Enginyeria Industrial de Barcelona, Universitat Politècnica de Catalunya, Barcelona, Spain

Francesc Rocadenbosch

Department of Signal Theory and Communications, Lidar Group, Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona, Universitat Politècnica de Catalunya, Barcelona, Spain

José M. Baldasano

Laboratory of Environmental Modeling, Escola Tècnica Superior d'Enginyeria Industrial de Barcelona, Universitat Politècnica de Catalunya, Barcelona, Spain

ABSTRACT

A cluster algorithm was implemented to analyze the backward trajectories arriving in the Barcelona area (BCN), located at the eastern coast of the Iberian Peninsula. Five years of 4-day kinematic back trajectories, computed with version 4 of the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT), were clustered and classified in groups of similar length and curvature. To describe better the tropospheric circulations at BCN, backward trajectories arriving at 5500, 3000, and 1500 m above sea level were analyzed. The main transport patterns are identified at 5500 m: westerly flows (49% of the total situations), northwesterly flows (17%), southwesterly flows (20%), and regional recirculations over Europe and the Mediterranean Sea (15%). An annual distribution of the transport patterns is described. During the summertime, moderate westerlies (19%) and southwesterlies (13%) in the middle troposphere, slow westerlies (11%) and southwesterlies (10%) at 3000 m, and regional recirculations (29%) at 1500 m characterize the BCN long-range transport. This general pattern varies during wintertime, with more westerlies at 5500 m and an increase of northern and northwestern situations at 1500 m. A large number of situations with decoupling between the lower and middle troposphere are observed when combining 5500- and 1500-m cluster results. Interannual variability is discussed, and the influence of the North Atlantic Oscillation phase is captured by winter average regime patterns.

1. Introduction

The Barcelona area (BCN) is located at the eastern coast of the Iberian Peninsula (IP) within the western Mediterranean basin (WMB). Several studies of synoptic climatology were developed for the IP during the past 30 yr to characterize the weather type and the typical synoptic situations affecting the region (e.g., Albentosa 1973; Font-Tullot 1983; Martín-Vide 1987, 1991; Sánchez 1993; Clavero et al. 1996; Capel-Molina 2000). Most of these works were based on the authors' knowledge of the local and regional meteorological behavior to develop distinct subjective classifications of flows. Martín-Vide (1987) developed a flow climatology defining 20 typical synoptic situations for the IP by analyzing a 30-yr series of weather maps (surface pressure and 500-hPa topography). At that time, the computational capabilities and resources were very limited relative to the past decade and nowadays. Several studies with more objective approaches are demanded (Martín-Vide 2002).

In order to reduce the subjectivity of these classifications and to improve the analysis of large meteorological datasets, cluster analysis appeared at the end of the 1980s as a valuable tool to be applied to meteorological data (Kalkstein et al. 1987). Moody (1986) and Moody and Galloway (1988) were the first to perform studies using cluster analysis with atmospheric trajectories. They applied clustering techniques to interpret trajectory and precipitation chemistry data. In the same way, several authors have analyzed the influence of atmospheric transport patterns on pollutants concentrations applying trajectory clusters (e.g., Moody and Sam-

Corresponding author address: José M. Baldasano, Laboratory of Environmental Modeling, ETSEIB, Universitat Politècnica de Catalunya, Av. Diagonal 647, Office 10.23, 08028 Barcelona, Spain.
E-mail: jose.baldasano@upc.es

TABLE 1. Number per year and percentage of available back trajectories.

Period	No.	Percent available
Jul 1997–Dec 1997	328	89.1
1998	711	97.4
1999	730	100
2000	714	97.5
2001	720	98.6
Jan 2002–Jun 2002	362	100
Total	3565	97.6

son 1989; Moody et al. 1995; Dorling et al. 1992a,b; Dorling and Davis 1995; Brankov et al. 1998; Avila and Alarcón 1999; Cape et al. 2000). In principle, the results of cluster analysis are similar to those of flow climatology, but the former technique is more objective and accounts for variations in transport speed and direction simultaneously, yielding clusters of trajectories that have a similar length and curvature (Stohl 1998). Harris and Kahl (1990) and Harris (1992) based a flow climatology for the Mauna Loa Observatory in Hiloa, Hawaii, and the South Pole, respectively, clustering a large number of back trajectories. For the IP, some efforts to develop studies of synoptic climatology with a more objective approach were done. Some of them clustered surface pressure and geopotential maps (Calvo 1993; Petisco and Martín 1995), others clustered precipitation patterns to obtain a synoptic classification (Ribalaygua and Borén 1995), and also some authors have tested the automatic synoptic classification of Jenkinson and Collison (1977) to the IP (Spellman 2000), detecting some weakness in the methodology (Martín-Vide 2001). Avila and Alarcón (1999) present one of the few works using cluster analysis with atmospheric trajectories for the IP, but it analyzes only days with precipitation to interpret precipitation chemistry data. The application of cluster analysis to a database of meteorological trajectories has resulted in a valuable approach to the description of long-range transport of air masses.

Herein, we describe the tropospheric flow patterns arriving at BCN by means of cluster technique and a 5-yr database (July 1997–June 2002) of atmospheric back trajectories. It represents an extended approach of this technique for the IP. Cluster analysis was used to group trajectories according to wind speed and direction in order to describe the main flows arriving at BCN and to identify the origin of the air masses affecting the region. The aim of the work is not to provide a climatology of the flows affecting BCN, but to summarize and quantify the flow characteristics of the region with the application of multivariate statistical techniques to complement the several studies of synoptic climatology and long-range transport developed for the IP.

2. Cluster analysis

Cluster analysis is a multivariate statistical technique designed to explore structure within a dataset (Ander-
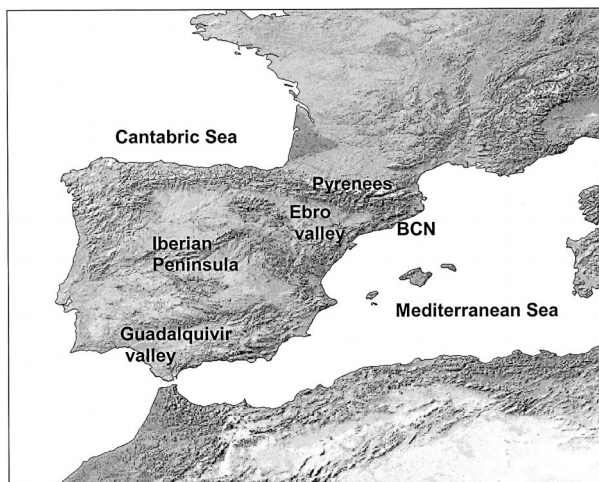

FIG. 1. The Iberian Peninsula and the Barcelona region (toponyms referred to in the text are indicated).

berg 1973; Everitt 1980). It is often described as an objective classification method, but this is not absolutely true because the selection of the clustering algorithm, the specifications of the distance measure, and the number of clusters used are subjective (Stohl 1998). While many different clustering algorithms exist, there is much variation in the computational requirement for interpretation of the data. The nonhierarchical clustering algorithm employed for this work, based on Dorling et al. (1992a) and Mattis (2001), is specially designed for large databases because of its relatively small computational requirements. Its main feature is that the optimum number of clusters follows from the algorithm itself and does not have to be assumed previously or evaluated with another technique (e.g., Harris and Kahl 1990; Moody et al. 1995).

The algorithm procedure is described in detail by Dorling et al. (1992a). The modifications introduced concern the initialization of the process. While Dorling et al. (1992a) generate a large number of "seed" trajectories, which cover the spread of the real trajectories used in the analysis, the methodology employed in this study is based on Mattis (2001), generating a large number of synthetic seed trajectories rather than real seed trajectories. The advantage of using synthetic trajectories to initialize the procedure relies on a faster convergence of the solution. The synthetic trajectories cover all of the spread of the back trajectories uniformly and lead to a faster resolution of the clustering procedure. The cluster process was first performed for all seeds, then every sixth seed trajectory was left out, starting with the first one; this process is repeated successively, following Dorling et al. (1992a). The measure of distance used to evaluate the root-mean-square deviation (rmsd) of a trajectory from its centroid was based on the Haversine formula of the great-circle distance between two points (Sinnott 1984), concerning only latitude and longitude variables [Eq. (1)],
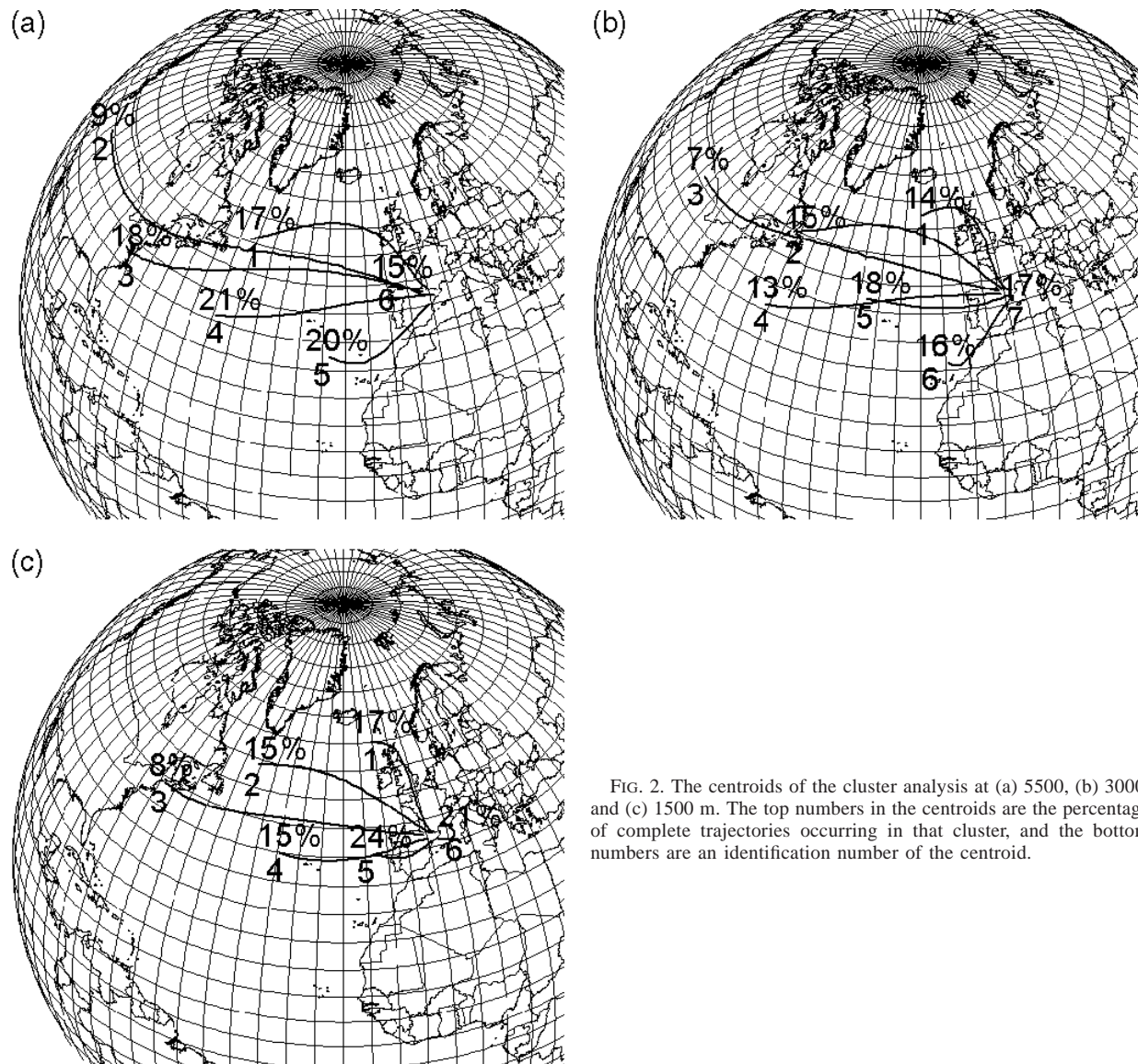
FIG. 2. The centroids of the cluster analysis at (a) 5500, (b) 3000, and (c) 1500 m. The top numbers in the centroids are the percentage of complete trajectories occurring in that cluster, and the bottom numbers are an identification number of the centroid.

$$rmsd = \sqrt{\frac{1}{N} \sum_{i=1}^{N} D_i^2}, \quad (1)$$

where D_i is the absolute distance between trajectories evaluated as the sum of the distances between 6-hourly coordinate points using the Haversine formula. The number of back trajectories analyzed is N.

The optimal number of clusters is obtained by plotting the total rmsd (sum of rmsds of each trajectory from its centroid) against the number of clusters. This graph presents a monotonic increase, where sudden breaks are evident at various stages of the clustering process. These are interpreted as the merging clusters of trajectories that significantly differ in terms of wind direction and speed, which they represent (Dorling et al. 1992a). In order to have a manageable number of clusters for further descriptions, the authors assumed that a 5% change

in total rmsd is significantly large and, thus, the number of clusters is retained. This threshold is adopted in other works as Dorling et al. (1992a) and Brankov et al. (1998).

3. Trajectory database

In order to describe the flows reaching the middle troposphere, the low free troposphere, and the upper-boundary layer over BCN, 4-day kinematic back trajectories arriving at three different heights [5500, 3000, and 1500 m above mean sea level (MSL)] were computed twice daily (0000 and 1200 UTC) for the period from July 1997 to June 2002. Because a trajectory is best used as an indication of the general airflow rather than the exact pathway of an air parcel, it is hoped that analysis of a large number of trajectories in a clima-
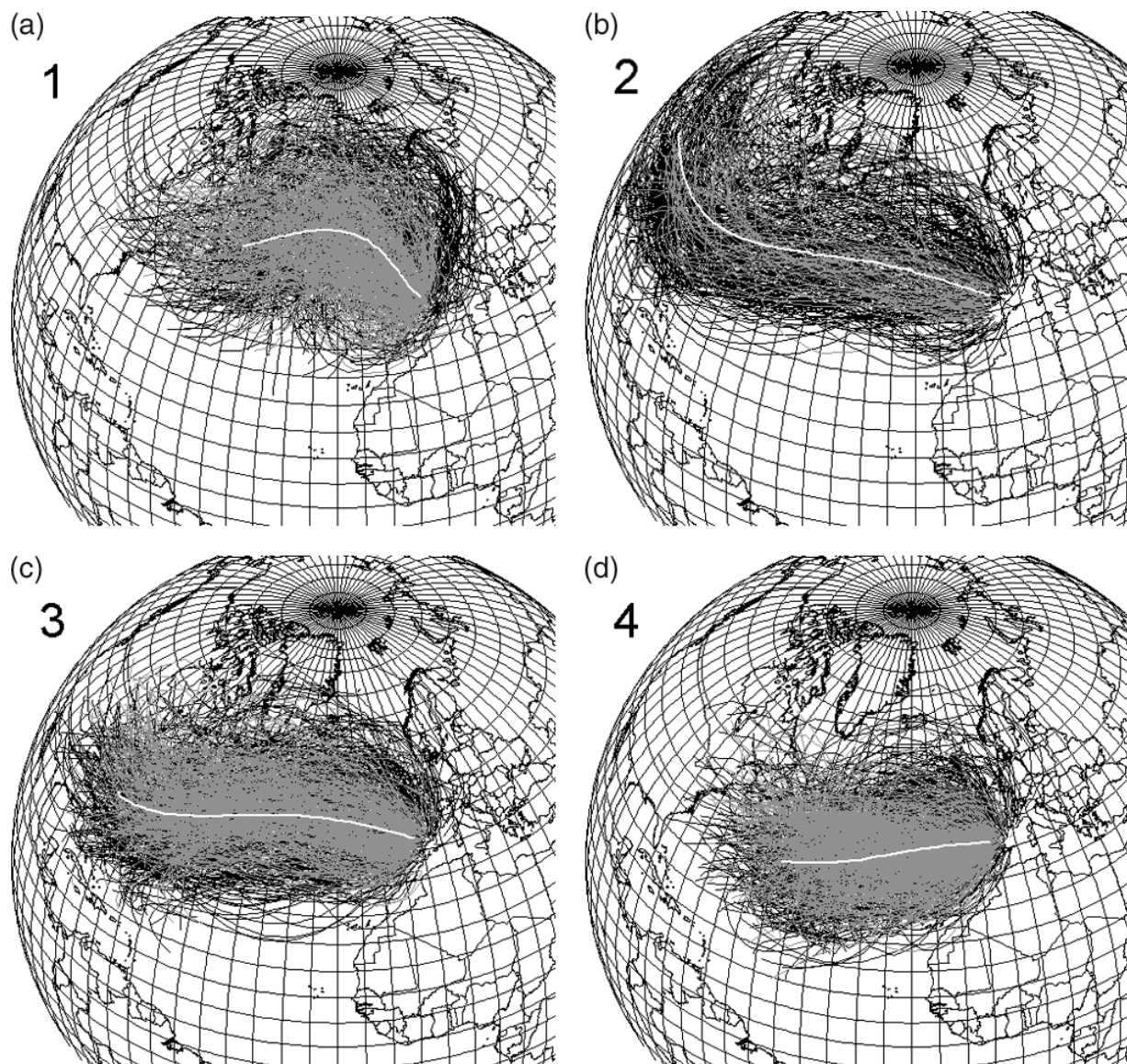
FIG. 3. (a)–(f) The six cluster membership plots for 4-day 5500-m back trajectories for the period from Jul 1997 to Jun 2002 (Black lines: winter back trajectories, gray lines: summer back trajectories, white line: centroid).

tological sense will reduce the effects of individual errors (Harris and Kahl 1990). Hence, a reasonable representation of the airflow to BCN is possible.

Kinematic back trajectories were calculated with version 4 of the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT) developed by the National Oceanic and Atmospheric Administration (NOAA)'s Air Resources Laboratory (ARL) (Draxler and Hess 1998; Draxler and Rolph 2003; Rolph 2003). The meteorological data used for the computation of the trajectories comes from the "FNL" archive maintained by ARL (available online at http://www.arl.noaa.gov/ss/transport/archives.html). The 6-hourly archive data come from The National Centers for Environmental Prediction (NCEP) Global Data Assimilation System

(GDAS; Kanamitsu 1989), which uses the spectral Medium-Range Forecast model (MRF) for the forecast. A summary of the trajectory database is given in Table 1.

Kinematic 3D trajectories were used following the suggestion of several authors about the greater accuracy of these trajectories in comparison with all of the other approaches (isentropic, isobaric) when accurate fields of vertical wind are available (e.g., Martin et al. 1990; Draxler 1996; Stohl and Seibert 1998). A recent review on computation and applications of trajectories was provided by Stohl (1998). The length of the back trajectories is restricted in many ways by the distances between source regions and the destination zone. The selection of 4-day back trajectories was done because there were enough representative of the long-range transport
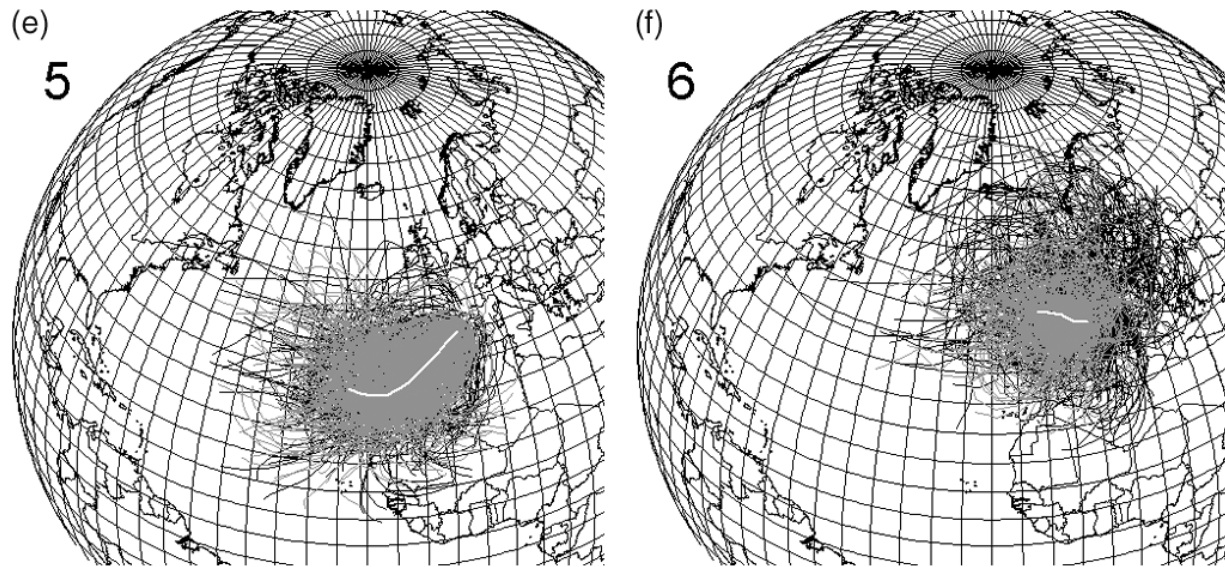
FIG. 3. (*Continued*)

to the IP, and its error remains controlled. Stohl (1998) affirmed that errors of 20% of the distance traveled seem to be typical for trajectories computed from analyzed wind fields.

4. Results and discussion

a. Synoptic features affecting flow to BCN

BCN is located at the eastern coast of the IP within the WMB (Fig. 1). The Mediterranean basin is usually wedged between two major mean synoptic pressure fea-

TABLE 2. RHTD from the average pattern at 1, 2, 3, and 4 days before the arrival time of back trajectories for the 5-yr cluster results in percentage.

	Clusters	RHTD $T = -1$ day (%)	RHTD $T = -2$ days (%)	RHTD $T = -3$ days (%)	RHTD $T = -4$ days (%)
5500 m	1	97	55	39	38
	2	47	24	17	18
	3	57	29	18	20
	4	64	34	25	28
	5	53	39	39	46
	6	253	231	196	139
3000 m	1	95	64	48	50
	2	75	38	24	30
	3	49	27	21	22
	4	51	28	20	24
	5	65	45	34	38
	6	56	41	40	48
	7	484	400	375	260
1500 m	1	86	60	48	51
	2	69	41	28	34
	3	49	30	24	27
	4	57	39	31	35
	5	100	80	70	71
	6	287	173	154	150

tures. In winter the Azores high pressure system is located over the Atlantic Ocean and the Siberian high pressure system is over Eastern Europe and Asia. In summer, the Azores high strengthens and remains centered over the open Atlantic, while the Siberian high gives way completely to the influence of the strong summer thermal low over southwestern Asia, which becomes the engine that drives the intense southwest monsoon over the Arabian Sea. The IP in winter is characterized by an increase of extratropical cyclone activity, and it is the wet season with the strongest winds. On the other hand, during the summer season the subtropical highs increase their influence and, because of subsidence and the resulting stable lapse rate, produce characteristically hot dry summers. The mean summer wind pattern is similar to that of winter but with much lower speeds (Elms et al. 1987).

The mountain ranges surrounding the Mediterranean Sea act as a sharp climate barrier protecting the Mediterranean basin from much of the more extreme continental weather conditions. Local winds within the Mediterranean area are generated and influenced by mountainous terrain surrounding the Mediterranean Sea. These mountains can funnel the wind flow into the basin, often creating higher wind speeds that can reach gale or storm force strengths. The major orographic features that influence the flows to BCN are the Pyrenees mountains and the Ebro Valley. Typical local winds affecting BCN are the tramontana (a northern wind), and the cierzo (a northwestern wind channeled through the Ebro Valley). The development of sea breeze and mountain–valley winds, and the typical Iberian thermal low are characteristic features of the region under summer situations. Although the mesoscale effects may not be captured by the analyses from which the trajectories are
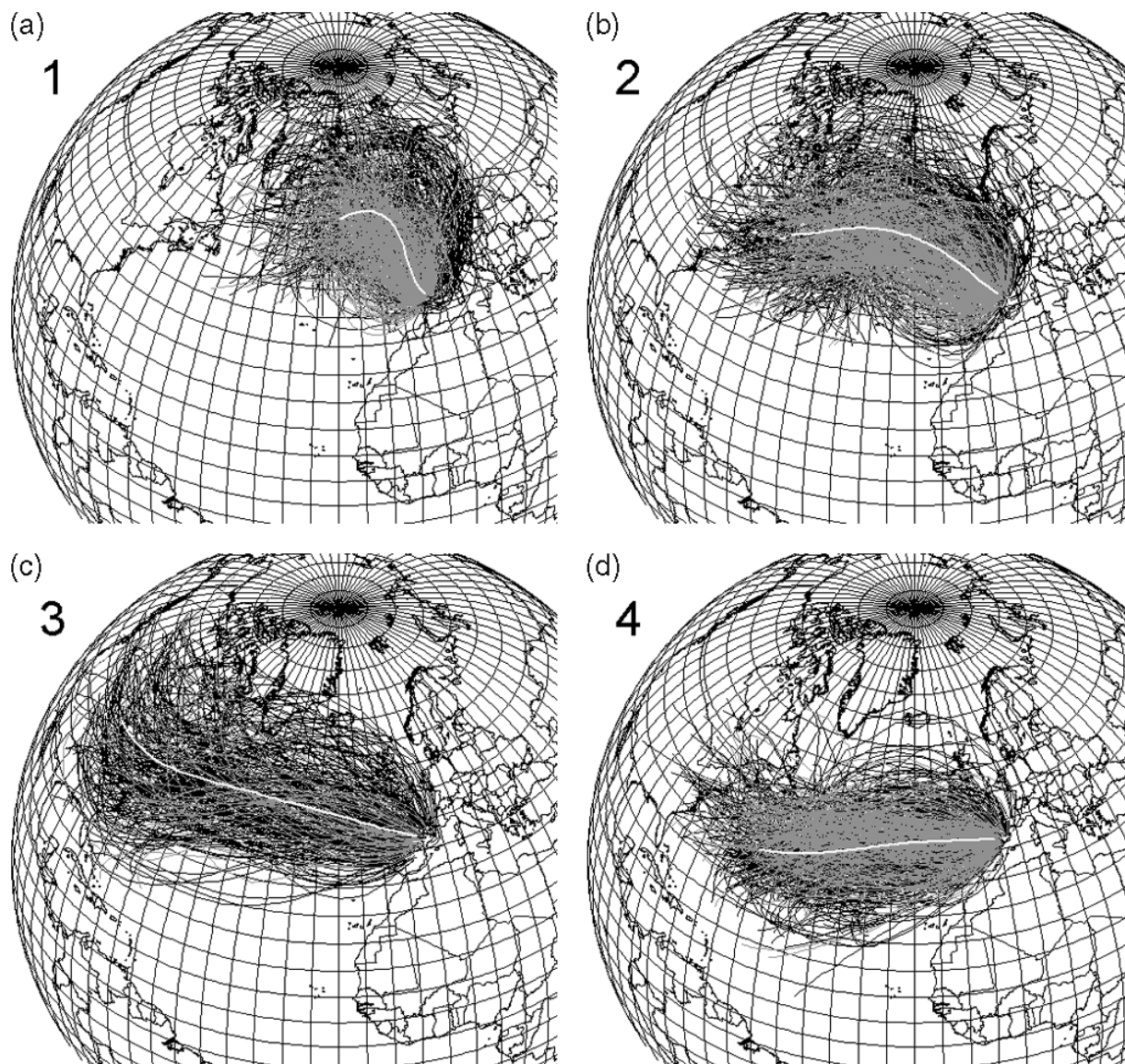
FIG. 4. (a)–(g) Same as Fig. 3, but for the seven clusters at 3000 m.

created, the back trajectories help to identify situations in which these mesoscale effects should develop.

b. 5500-m transport regime

The cluster analysis has established six principal groups of trajectories arriving at 5500 m at BCN. The cluster-mean trajectories, hereinafter referred to as centroids, are shown in Fig. 2a. The dominant 5500-m transport regimes from July 1997 to June 2002 include northwesterly flows (cluster 1), westerly flows (clusters 2, 3, and 4), southwesterly flows (cluster 5), and regional recirculations over Europe (cluster 6). Figure 3 displays the composite of all trajectories within each cluster for the whole period. Substantial variability within individual clusters is evident. These plots give a qualitative

feeling of the variability within each cluster as noted by Harris and Kahl (1990).

The centroids represent the averaged flow of the complete back-trajectory group. Therefore, a general description of the main transport features of a cluster can be done with its centroid. The relative horizontal transport deviation (RHTD) of every cluster has been calculated as a measure of how representative the centroids are. Table 2 presents these values. RHTD was calculated as the absolute horizontal transport deviation (e.g., Stohl 1998) divided by the length of the centroid (the length of the curved trajectory, not the straight line between the starting and ending points) at 1, 2, 3, and 4 days before the arrival time of the trajectory. From the results, some care must be taken with slow-moving centroids (e.g., Fig. 3f). This centroid trajectory can be mislead-
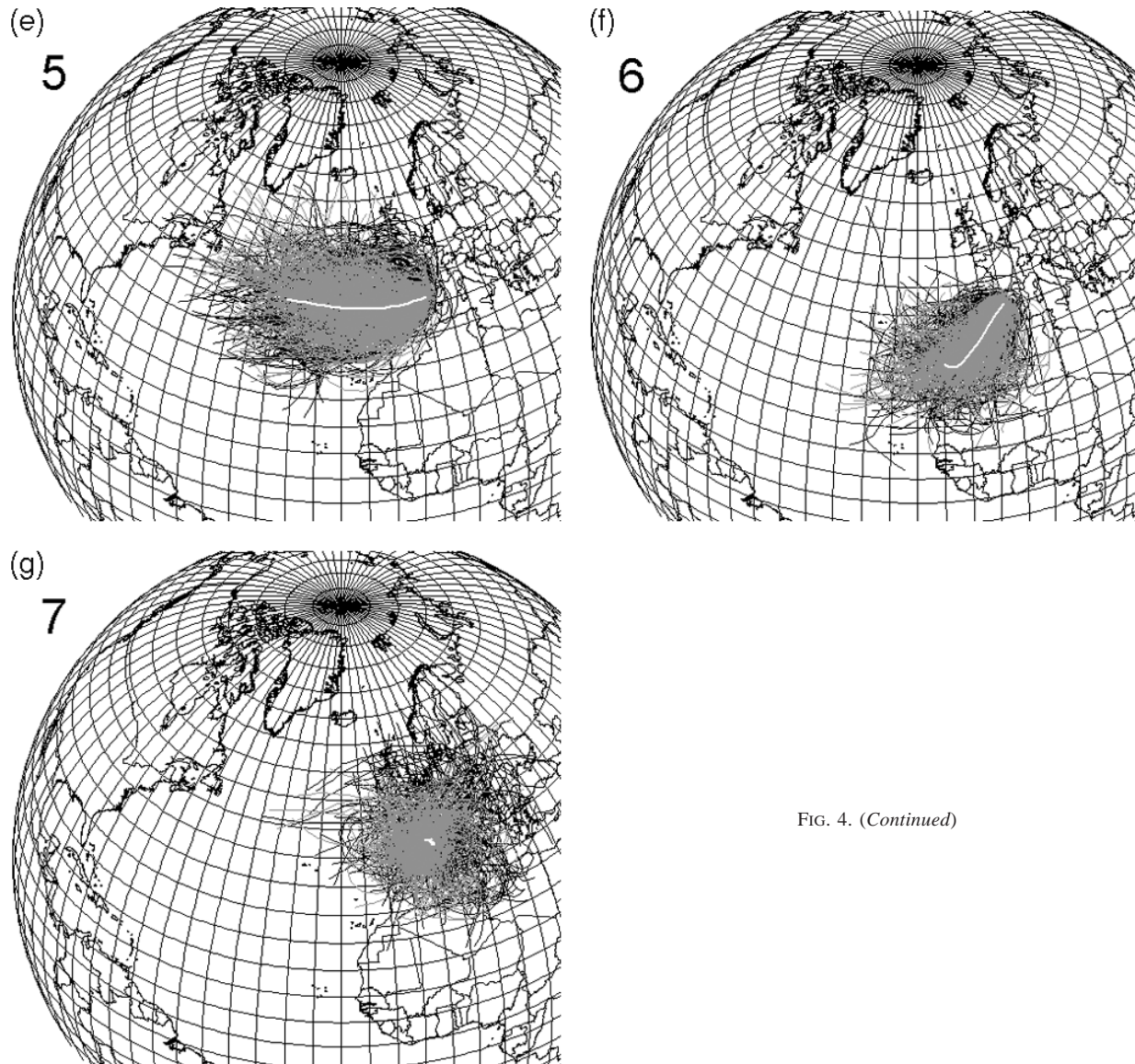
(e)

5

(f)

6

(g)

7

FIG. 4. (*Continued*)

ing, because the variability within the cluster may be much greater than the length of the centroid trajectory, with RHTD greater than 150%. Thus, the interpretation of short, slow-moving cluster-means requires consideration of the corresponding membership plot (Harris and Kahl 1990). These flows are mainly regional recirculations over a relatively limited region. Air masses may recirculate over the zone during several days under anticyclonic situations, especially during summer (Millán et al. 1997). When the flow presents a marked advection component, the RHTD presents lower values. The RHTD usually remains below 50%, suggesting that the centroids are representative of their clusters as an average transport pattern.

Table 3 summarizes the cluster results based on the 5 yr of back trajectories. Zonal flows account for 48% of the data. These flows are grouped into fast westerly

flows (cluster 2) with a 9% of occurrence, moderate westerly flows (cluster 3) representing 18%, and slow westerly flows (cluster 4) accounting for a total of 21%. Transport of air masses from North America is produced under this transport regime. Cluster 2 is remarkable, because it includes very long trajectories originating 4 days back over the Pacific Ocean. Also, cluster 3, which accounts for moderate westerly winds, may be responsible for long-range transport of American continental air masses to the Mediterranean basin. Slow westerly flows present a more southern direction, advecting air masses from the central Atlantic Ocean.

Another transport regime that accounts for 17% of the data is the northwesterly flows represented by centroid 1. This cluster is made up of northwestern, northern, and some northeastern flows, the latter occurring especially under winter situations. The length of the
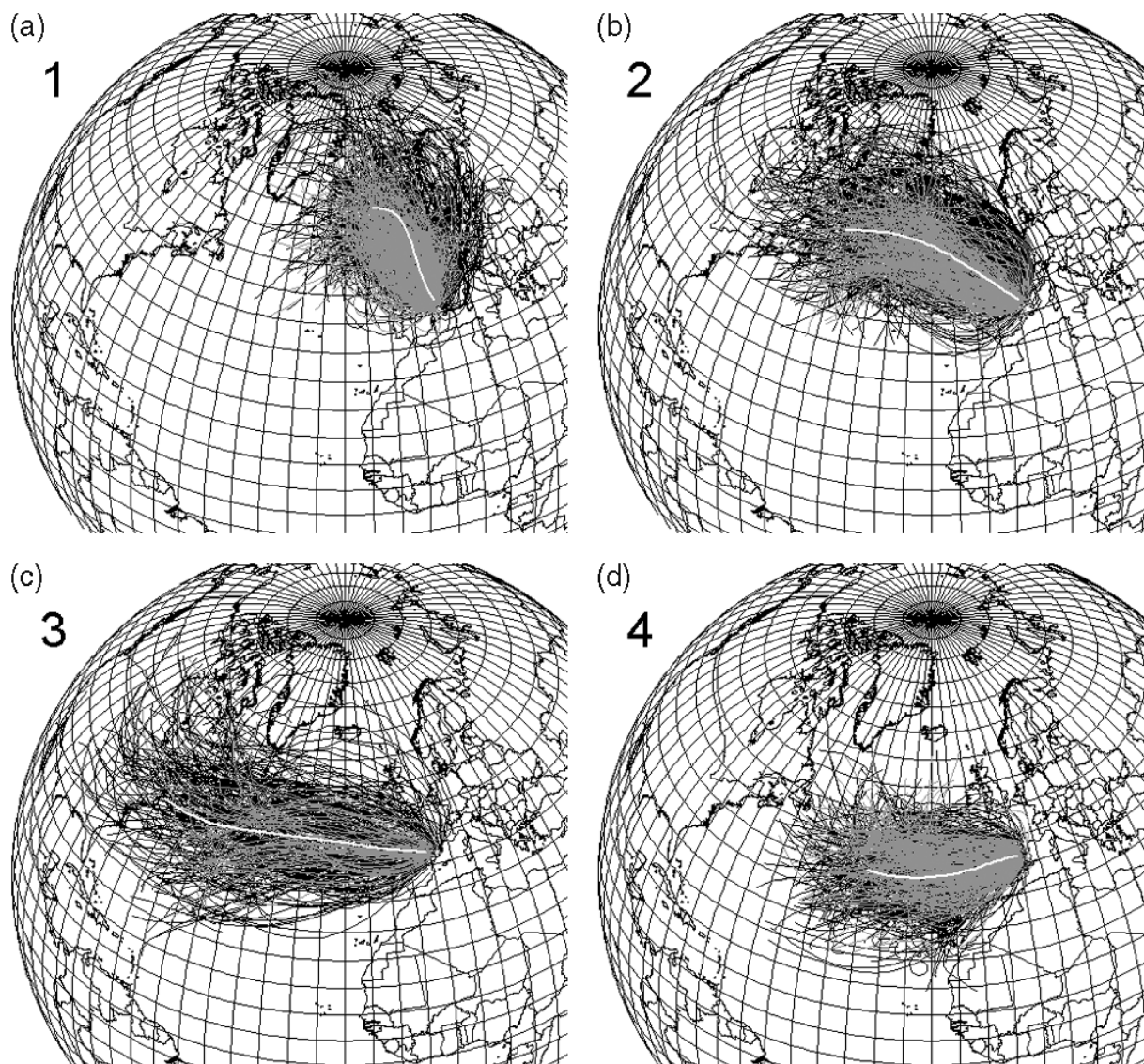
FIG. 5. (a)–(f) Same as Fig. 3, but for the six clusters at 1500 m.

trajectories is enough to transport air masses from eastern Canada. The mean shape of the cluster presents an anticyclonic curvature, advecting polar maritime air, arctic maritime air, and in some cases polar continental air masses to BCN.

Southwesterlies represent 20% of the cases (cluster 5). Slow-blowing winds from middle latitudes advect warm air masses with the possible transport of Saharan dust to the Mediterranean and central Europe. These situations occur predominantly during the summer with a 13% frequency of occurrence between April and September. The Saharan dust air masses reach the northeastern IP predominantly by depressions located to the west or southwest of the IP, when the North African anticyclone shifts to the east or southeast of the IP, or by the combination of both cyclone and anticyclone systems (Querol et al. 2001; Rodríguez et al. 2001). The

short lengths of the trajectories are indicative of possible recirculations over southwestern Europe, or circulations around cyclogenic zones or high pressure centers.

Last, 15% of the cases are included within a regional recirculation cluster. During summer the quasi-persistent influence of the Azores high over the region, and in winter the disposition of the European continental anticyclone and its influence over the Mediterranean air basin, can be reflected at middle troposphere with a low baric gradient situation.

c. 3000- and 1500-m transport regime

The flows within the low free troposphere and the upper boundary layer were also analyzed. Thus, back trajectories arriving at 3000 and 1500 m were computed and clustered for the period from July 1997 to June
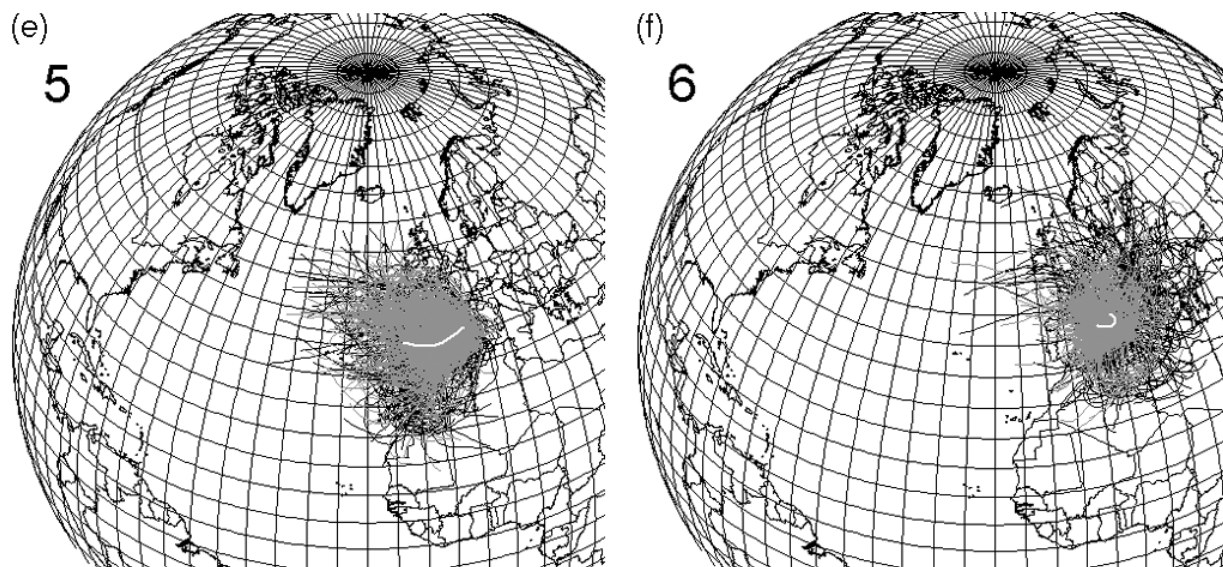
FIG. 5. (*Continued*)

2002. Seven clusters were obtained for the 3000-m trajectories and six clusters for the 1500-m trajectories. Figures 2b and 2c show the centroids for both levels, Figs. 4 and 5 depict their cluster-membership plot, and Table 3 presents a summary of the results. Some differences appear when comparing both results. Southwesterlies are clearly defined in the 3000-m results in cluster 6 (Fig. 4f), but they have a lower occurrence at 1500 m without a specific cluster for them; they are included in cluster 5 (Fig. 5e). Southerly flows represent one of the biggest clusters in the free troposphere.

The 1500-m back-trajectory results point out an elevated occurrence of stagnant pressure systems over the region that produce several days of recirculations of air masses. This is clearly featured in cluster 5 (western recirculations) and cluster 6 (eastern recirculations) presented in Figs. 5e and 5f, which account for a total of 45% of the 1500-m back trajectories. Western recirculations include very slow westerly flows, some recirculations over the Iberian Peninsula, and some southern flows advecting warm air from the African continent to BCN. The cluster of eastern recirculations contains trajectories recirculating over the Mediterranean Sea and Eastern Europe. In summer, regional recirculations are frequent at low levels (16% and 13% for western and eastern recirculations, respectively) because of the influence of the Azores high pressure system, which produces several days of similar low baric gradient situations at low levels, while in the middle troposphere weak westerlies and northwesterlies are common. Hence, the development of mesoscale phenomena, induced by the particular topography of the region, is dominant. These phenomena are mainly sea breezes, up-slope and down-slope winds, and valley-channeled winds, and the development of a large mesoscale circulation with marked diurnal cycles: the Iberian thermal low with its com-

pensatory return flows over the seas (Millán et al. 1997). The strong heating of the surface air over the Spanish plateau and the Guadalquivir Valley causes low-level convergence and ascent of the air masses, contributing to the development of the Iberian thermal low. Some of these situations are associated with local episodes of air pollution in the region that result in higher levels of ozone and an increase of particulate matter within the boundary layer in BCN during summer (e.g., Millán et al. 1997; Toll and Baldasano 2000; Barros et al. 2003). Also, these recirculations are associated with a multi-layer arrangement of aerosols observed over BCN under typical summertime situations in the absence of large-scale forcing (Soriano et al. 2001). In winter, the influence of the European anticyclone can also produce several days of stagnant high pressure over BCN with a poor development of the boundary layer—less than 500 m MSL in some cases (Sicard et al. 2003). Under this situation, recirculations of air masses over BCN are common, but with less intensity than during summer.

The cluster accounting for advections of northern air masses, some fast recirculations over the central part of Western Europe, and several continental northeasterly flows (Figs. 4a and 5a) represents 14% of the 3000-m and 17% of the 1500-m trajectory database. Those circulations present a major occurrence in winter, especially at low levels.

Another important cluster groups northwesterly flows, representing 15% of the 3000- and 1500-m back trajectories. It accounts for fast flows of polar maritime air masses; some of them start in eastern Canada, and in the last days travel over Western Europe.

Westerly flows are grouped into three clusters at 3000 m (Figs. 4c–e): fast westerly flows, westerly flows, and slow westerlies. Under this situation, transport of air masses from the United States and Canada to BCN, the

TABLE 3. Summary of 5500-, 3000-, and 1500-m 5-yr seasonal, and monthly back-trajectory cluster results (summer: Apr–Sep; winter: Oct–Mar; percentage values are calculated with respect to the 3565 back trajectories used in the cluster analysis).

Arrival height	5500 m														
Cluster	Tot (%)	Summer (%)	Winter (%)	Jan (%)	Feb (%)	Mar (%)	Apr (%)	May (%)	Jun (%)	Jul (%)	Aug (%)	Sep (%)	Oct (%)	Nov (%)	Dec (%)
Northwesterlies (NW)	17	8	9	1.3	1.4	1.3	1.8	1.0	1.1	1.5	1.4	1.7	0.9	1.9	1.7
Fast westerlies (W)	9	2	7	1.4	0.9	0.8	0.7	0.4	0.2	0.2	0.1	0.4	1.1	1.5	1.3
Westerlies (W)	18	8	10	1.8	1.5	1.8	2.0	1.0	1.2	1.4	1.1	1.2	1.4	1.5	1.8
Slow westerlies (W)	21	11	10	1.7	1.6	1.8	2.0	1.7	2.1	1.9	1.5	1.4	1.9	1.3	2.1
Southwesterlies (SW)	20	13	7	1.1	0.9	1.8	1.0	2.4	2.2	2.4	2.9	1.9	1.7	0.9	0.8
Regional recirculations (R)	15	8	7	1.4	1.3	1.0	0.8	2.0	1.3	0.9	1.4	1.3	1.4	1.2	1.0

Arrival height	3000 m														
Cluster	Tot (%)	Summer (%)	Winter (%)	Jan (%)	Feb (%)	Mar (%)	Apr (%)	May (%)	Jun (%)	Jul (%)	Aug (%)	Sep (%)	Oct (%)	Nov (%)	Dec (%)
Northerlies (N)	14	7	7	1.1	0.9	1.3	1.4	0.9	0.7	1.3	1.0	1.2	0.9	1.7	1.6
Northwesterlies (NW)	15	7	8	1.2	1.3	1.6	1.6	0.6	1.1	1.2	0.8	1.3	1.1	1.8	1.4
Fast westerlies (W)	7	1	6	1.2	0.8	0.9	0.6	0.2	0.2	0.1	0.0	0.2	1.0	1.2	0.8
Westerlies (W)	13	5	8	1.3	1.5	1.2	1.7	1.1	0.7	0.7	0.4	0.6	0.8	1.0	2.0
Slow westerlies (W)	18	11	7	1.1	1.3	1.4	1.8	1.3	2.1	2.0	2.2	1.5	0.9	1.0	1.4
Southwesterlies (SW)	16	10	6	1.0	0.7	1.3	0.5	2.0	1.9	1.7	2.3	1.7	1.4	0.8	0.7
Regional recirculations (R)	17	10	7	1.8	1.3	1.0	1.0	2.4	1.6	1.2	1.7	1.5	1.2	1.3	1.0

Arrival height	1500 m														
Cluster	Tot (%)	Summer (%)	Winter (%)	Jan (%)	Feb (%)	Mar (%)	Apr (%)	May (%)	Jun (%)	Jul (%)	Aug (%)	Sep (%)	Oct (%)	Nov (%)	Dec (%)
Northerlies (N)	17	7	9	1.0	1.6	1.9	2.0	1.4	0.8	1.6	0.6	1.1	1.3	1.9	1.8
Northwesterlies (NW)	15	6	9	1.1	1.5	1.5	1.6	0.6	1.0	0.9	0.4	1.3	1.3	2.4	1.4
Fast westerlies (W)	8	2	6	1.5	1.0	1.1	0.9	0.3	0.1	0.2	0.1	0.2	0.8	0.7	1.1
Westerlies (W)	15	6	9	1.4	1.5	1.6	1.5	1.2	1.0	0.7	0.7	0.8	1.4	1.1	2.1
Western recirculations (wR)	24	16	8	1.5	1.1	1.3	1.5	2.4	2.8	2.7	3.8	2.2	2.2	1.3	1.2
Eastern recirculations (eR)	21	13	8	2.2	1.1	1.3	0.9	2.7	2.4	2.2	2.6	2.1	1.2	1.1	1.2

WMB, and south Europe could occur. At 1500 m, westerlies are grouped into fast westerlies and westerlies (Figs. 5c–d). These regimes are more frequent during wintertime, representing 6% and 9%, respectively.

Results at 5500 and 1500 m have been combined in order to analyze the decoupling of the lower from the middle troposphere in the IP. The process followed consisted in considering a situation defined by their back trajectories at 5500 and 1500 m. Depending on where

TABLE 4. Frequencies of situations combining 5500- and 1500-m cluster results for the 5-yr analysis (the frequencies are calculated with respect to the 3565 back trajectories used in the cluster analysis; nomenclature of cluster situation detailed in Table 3, e.g., N: northerlies).

Situation 5500–1500-m cluster	No. of days	Frequency (%)	Situation 5500–1500-m cluster	No. of days	Frequency (%)
W–W	697	20	R–N	124	3
W–NW	366	10	R–wR	117	3
W–wR	323	9	SW–W	84	2
SW–wR	319	9	NW–eR	83	2
R–eR	283	8	NW–wR	78	2
SW–eR	260	7	NW–W	52	1
NW–N	251	7	SW–N	28	1
W–N	180	5	SW–NW	22	1
W–eR	139	4	R–NW	15	<1
NW–NW	133	4	R–W	11	<1

both trajectories were grouped in the cluster procedure, the situation is identified by the cluster group at 5500 and 1500 m [e.g., westerlies at 5500 m (W) and regional recirculations at 1500 m (R)]. Westerlies in 5500- and 1500-m cluster results were grouped together to reduce the number of situations obtained; thus, fast westerlies, westerlies, and slow westerlies are joined together as westerly flows at 5500 m, and fast westerlies and westerlies form a single westerly group at 1500 m. Hence, four clusters were considered at 5500 m and, in the same way, five at 1500 m. Table 4 presents the frequencies of the average situations formed when combining both levels. These results show important characteristics of the region. Coupled westerlies between both levels represent 20% of the situations, a low frequency in comparison with more northerly regions. Most of the situations present a decoupling of the lower from the middle troposphere. Situations with westerlies at middle levels and western recirculations at low levels represent 9% of the cases. The situations with a low baric gradient at low and middle levels represent 8% [R–eastern recirculations (eR)] and 3% [R–western recirculations (wR)] of the 5-yr-analyzed cases. The IP is characterized by several situations with decoupling between the lower and middle troposphere, and a large number of low baric gradient cases.
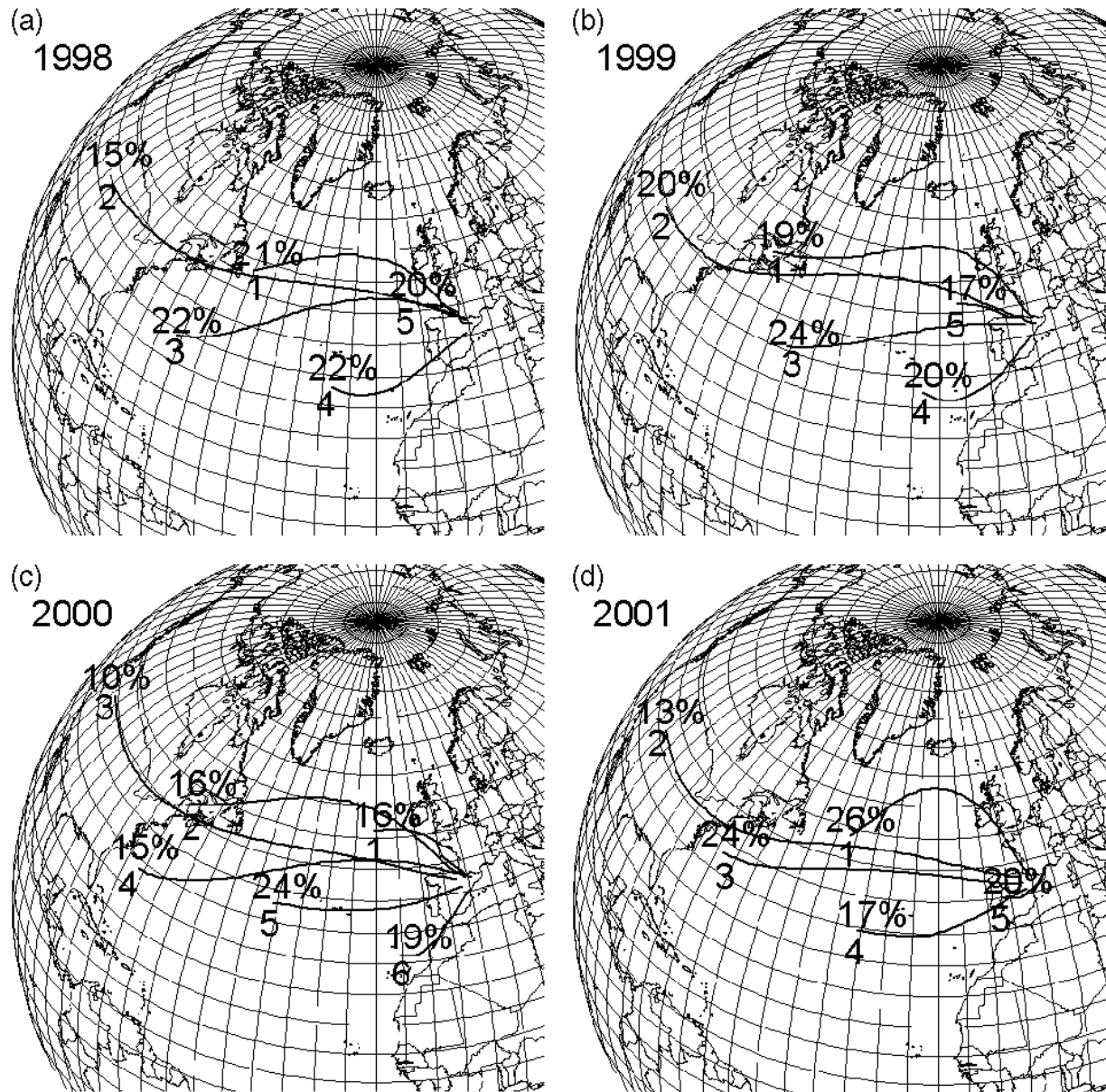
FIG. 6. Centroids plots for 4-day 5500-m back trajectories for the years (a) 1998, (b) 1999, (c) 2000, and (d) 2001. The top numbers in the centroids are the percent of complete trajectories occurring in that cluster, and the bottom numbers are an identification number of the centroid.

d. Year-to-year variability and monthly distribution

Cluster results may show variability from one year to the next. This makes sense due to the large interannual variability of the meteorological situations around the climatological behavior. The cluster algorithm was applied to the 1998, 1999, 2000, and 2001 databases individually. Cluster-mean plots for each year are shown in Fig. 6 in order to show the year-to-year variability in the 5500-m transport patterns.

Low variability is observed between 1998 and 1999. Only cluster 3 presents faster winds and a more northwestern arrival to BCN for 1998. Results of the year 2000 present one more cluster than the other cases, but show a similar behavior as that of 1998 and 1999. Year 2001 presents more direct westerly flows with a marked east–west direction to BCN in comparison with the other years. In 1998, 1999, and 2000 the northern centroids present a smoother northern component than in 2001, but have more marked southerly trajectories. Regional circulations over BCN, the IP, and the WMB are grouped in 1998, 1999, and 2001, with an occurrence of about 17%–20%. The analog recirculations for 2000 are grouped in cluster 1, presenting a longer path, with their spatial distribution displaced to the northwest.
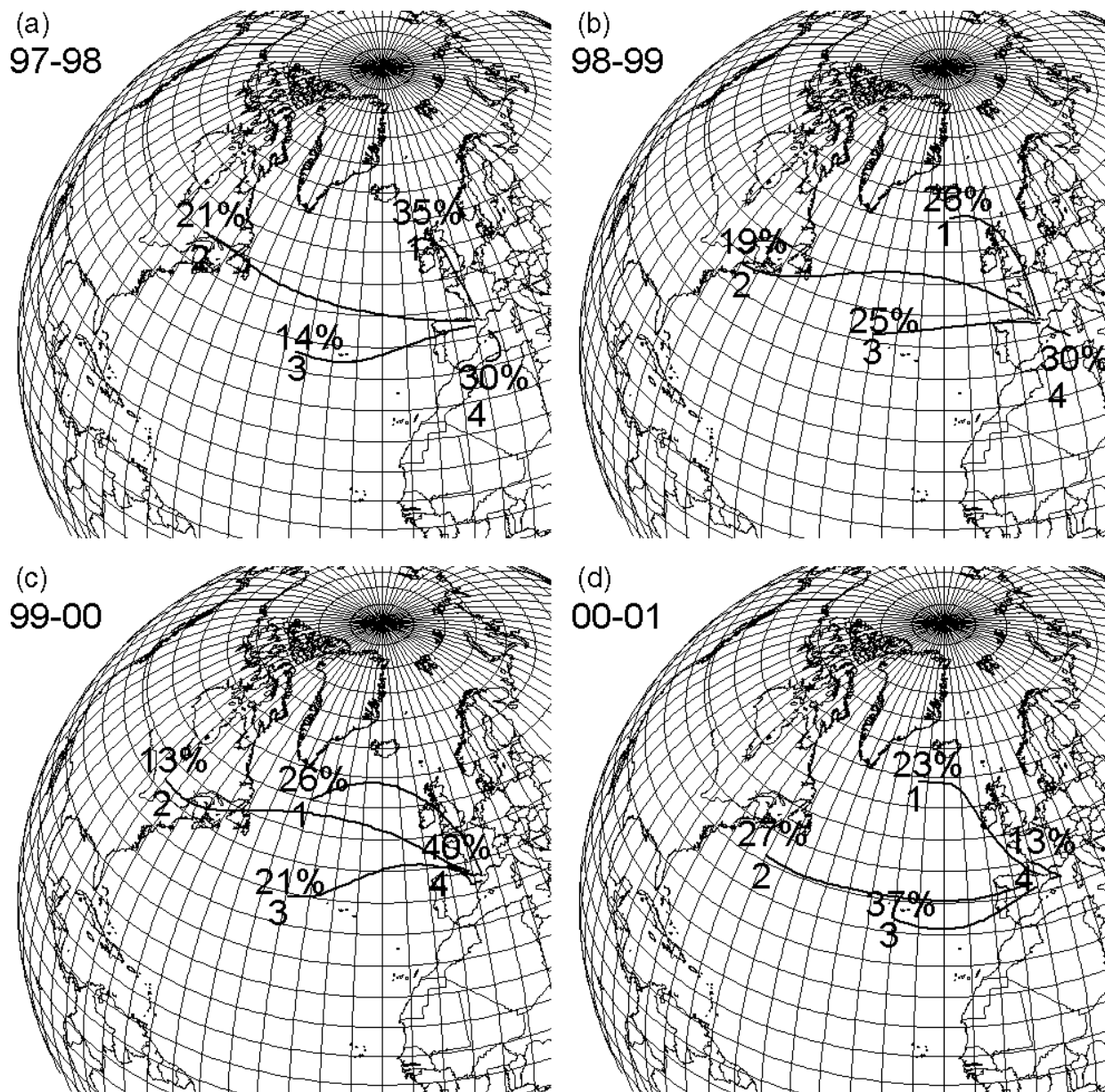
FIG. 7. Centroids plots for winter trajectories (Dec–Mar) of (a) 1997/98, (b) 1998/99, (c) 1999/2000, (d) 2000/01, and (e) 2001/02. The top numbers in the centroids are the percent of complete trajectories occurring in that cluster, and the bottom numbers are an identification number of the centroid.

Some of the differences could be in part explained by the influence of the North Atlantic Oscillation (NAO), which consists of a north–south dipole of pressure anomalies, with one center located over Iceland and the other center of opposite sign, spanning the central latitudes of the North Atlantic between 35° and 40°N. Years 1997, 1998, 1999, and 2000 were characterized by a positive phase of the NAO with an increase trend, followed by a reversal of the NAO in 2001, and again a positive phase in 2002 (Hurrell et al. 2003). The positive phase is associated with stronger-than-average westerly winds crossing the Atlantic Ocean on a more northerly track, and the negative phase is associated with weaker westerly winds crossing on a more west–east pathway. Yearly back-trajectory patterns do not clearly respond to the reversal of the NAO phase, but with the winter patterns this feature is well captured. The cluster process was applied to the winters of the studied period separately. Figure 7 shows the 1500-m winter back trajectory (December–March) cluster results for the periods of 1997/98, 1998/99, 1999/2000, 2000/01, and 2001/02. For the reversal of the NAO phase (winter of 2000/01) the westerly average pattern circulates at lower latitudes, and it pre-
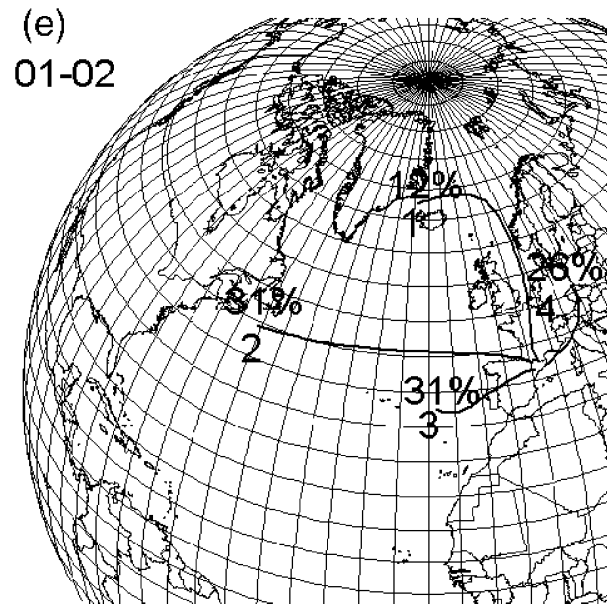
FIG. 7. (*Continued*)

sents slower winds in comparison with the other winters. Also, northerly flows have a more zonal arrival to BCN. This behavior is also observed at 5500 m (not shown).

The monthly distribution of the results explains the annual transport pattern characteristics in higher detail. Figure 8 and Table 3 present the monthly occurrence of the 5-yr database clusters for the three arrival heights. All of the percentage values are calculated with respect to the 3565 back trajectories of the analysis. A clear seasonal pattern is observed, with marked westerly and northerly regimes during winter, a transition during spring to more stagnant situations typical of summertime, and, finally, a return during autumn to the winter pattern. This evolution is observed in all three levels, but summer stagnant situations are more usual at low levels. Again, the decoupling between the lower and middle troposphere is well captured by the monthly distribution. Regional recirculations pattern at 1500 m occurs frequently during summer, having more than one-half of all trajectories (western and eastern recirculations). In the middle troposphere these situations are less common and present an annual regular distribution. This low occurrence is compensated by westerly flows, which are frequent at 5500 m. Fast westerlies are more important in the middle troposphere than at low levels. These flows are more usual in wintertime; November months present 1.5% of these situations. Northerly flows at 3000 m occur throughout the year at around 1% month^{-1}. November and December are the months with a higher occurrence of intrusions of northerly air masses, followed by April. In summer, an important flow regime is the southwesterly advection at middle troposphere. August months present 2.9% of these situations. This

flow is also important at 3000 m with a lower occurrence in comparison with 5500 m, but it represents around 1.5% month^{-1}, especially during the summertime.

5. Conclusions

A database of 4-day back trajectories for BCN computed twice daily at 5500-, 3000-, and 1500-m levels for the 5-yr period from July 1997 to June 2002 was used to describe the flow transport patterns over BCN. Cluster results have shown a marked zonal component in the average long-range transport patterns to BCN with three distinctive westerly groups (fast westerlies, westerlies, and slow westerlies) representing 48% at 5500 m, and 38% at 3000 m of the total analyzed situations. However, an elevated occurrence of regional recirculations is also observed, especially at low levels representing 45% of the situations at 1500 m (western recirculations and eastern recirculations). Important decoupling of the lower from the middle troposphere is observed, especially during summertime, as a distinctive characteristic of the region in comparison with more northern zones. Twenty different situations result when combining the 5500- and 1500-m cluster results, and 20% of the situations are characterized by westerlies at 5500 and 1500 m. It is noticeable that there is an elevated occurrence of situations with a low baric gradient at low levels and a more marked situation aloft [e.g., W–wR, southwesterlies (SW)–wR, W–eR].

Low yearly variability is observed between 1998 and 1999 average transport patterns, and more differences arise with 2000 and 2001 results. Part of the annual variability may be attributed to the NAO. The reversal of the NAO (2000/01) is not clearly captured by the
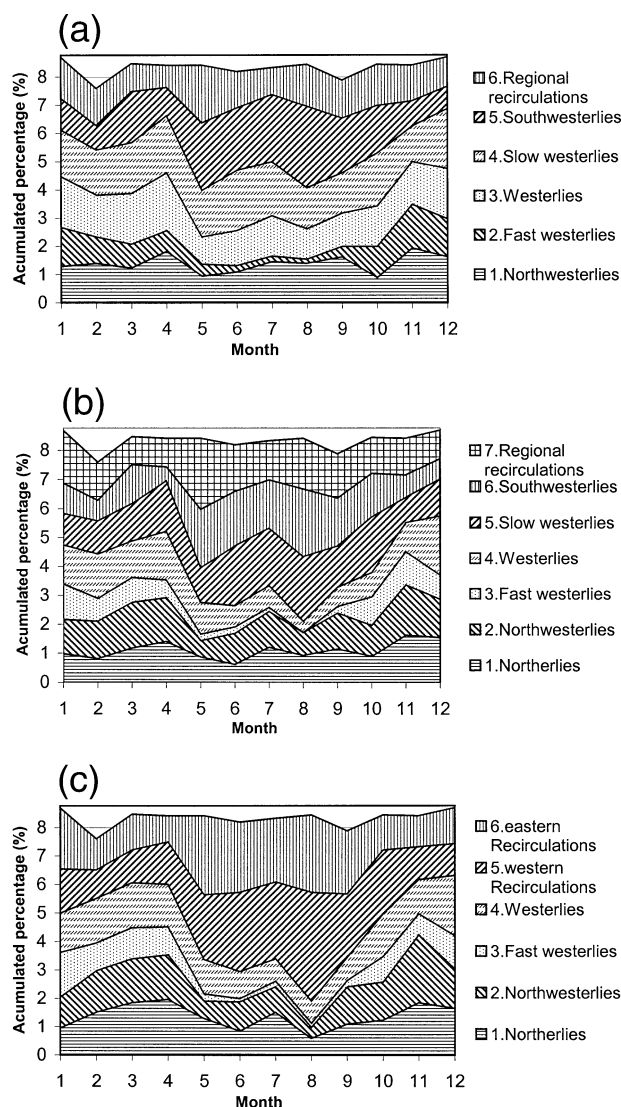
FIG. 8. Monthly distribution of clusters at (a) 5500, (b) 3000, and (c) 1500 m in percentage (percentage values are calculated with respect to the 3565 back trajectories used in the cluster analysis).

annual average transport patterns, but the winter patterns visibly reflect the reversal of the phase with westerlies arriving to BCN with a more marked east–west direction during the negative phase of the 2000/01 winter.

The present work summarizes the general flow characteristics of BCN with the application of multivariate statistical techniques to complement the studies of synoptic climatology and long-range transport developed for the IP. It contributes a quantitative approach to the description of the flows affecting BCN and IP, and provides a new approach to dynamic climatology of the IP that has a more objective methodology.

Some applications of the results may be related to describe the influence of atmospheric transport patterns on pollutant concentrations in BCN. Several situations with low pressure gradients in the region are associated

with development of photochemical air pollution episodes. Also, some authors have analyzed the long-range transport of pollutants from North America to Europe (e.g., Stohl and Trickl 1999). Similarly, episodes of the Saharan dust intrusions are usually observed during the summertime.

REFERENCES

Albentosa, L. M., 1973: Los climas de Cataluña: Estudio de climatología dinámica (Climates of Catalonia: Dynamic climatology study). Ph.D. dissertation summary, Universitat de Barcelona, 32 pp.

Anderberg, M. R., 1973: *Cluster Analysis for Applications.* Academic Press, 359 pp.

Avila, A., and M. Alarcón, 1999: Relationship between precipitation chemistry and meteorological situations at a rural site in NE Spain. *Atmos. Environ., 33,* 1663–1677.

Barros, N., L. Toll, C. Soriano, P. Jiménez, C. Borrego, and J. M. Baldasano, 2003: Urban photochemical pollution in the Iberian Peninsula: Lisbon and Barcelona airsheds. *J. Air Waste Manage. Assoc., 53,* 347–359.

Brankov, E., S. T. Rao, and P. S. Porter, 1998: A trajectory-clustering-correlation methodology for examining the long-range transport of air pollutants. *Atmos. Environ., 32,* 1525–1534.

Calvo, J., 1993: Clasificación de flujos en 500 hPa sobre la Península Ibérica (Flow classification at 500 hPa for the Iberian Peninsula). Service of Numerical Weather Prediction, Instituto Nacional de Meteorología, Tech. Note 32, 65 pp.

Cape, J. N., J. Methven, and L. E. Hudson, 2000: The use of trajectory cluster analysis to interpret trace gas measurements at Mace Head, Ireland. *Atmos. Environ., 34,* 3651–3663.

Capel-Molina, J. J., 2000: *El Clima de la Península Ibérica (The Iberian Peninsula Climate).* Ariel, 282 pp.

Clavero, P. L., J. Martín-Vide, and J. M. Raso, 1996: *Atles Climàtic de Catalunya: Termopluviometria (Climatological Atlas of Catalonia: Thermopluviometry).* Institut Cartogràfic de Catalunya, 41 pp.

Dorling, S. R., and T. D. Davies, 1995: Extending cluster analysis—synoptic meteorology links to characterise chemical climates at six northwest European monitoring stations. *Atmos. Environ., 29,* 145–167.

——, ——, and C. E. Pierce, 1992a: Cluster analysis: A technique for estimating the synoptic meteorological controls on air and precipitation chemistry—Method and applications. *Atmos. Environ., 26,* 2575–2581.

——, ——, and ——, 1992b: Cluster analysis: A technique for estimating the synoptic meteorological controls on air and precipitation chemistry—Results from Eskdalemuir, south Scotland. *Atmos. Environ., 26,* 2583–2602.

Draxler, R. R., 1996: Boundary layer isentropic and kinematic tra-

jectories during the August 1993 North Atlantic Regional Experiment Intensive. *J. Geophys. Res.,* **101,** 29 255–29 268.

——, and G. D. Hess, 1998: An overview of the HYSPLIT_4 modelling system for trajectories, dispersion, and deposition. *Aust. Meteor. Mag.,* **47,** 295–308.

——, and G. D. Rolph, cited 2003: NOAA ARL HYSPLIT Model. NOAA Air Resources Laboratory. [Available online at http://www.arl.noaa.gov/ready/hysplit4.html.]

Elms, J. D., and Coauthors, 1987: U.S. Navy Climatic Study of the Mediterranean Sea. U.S. Navy Climate Study NAVAIR 50-1C-457, 342 pp.

Everitt, B., 1980: *Cluster Analysis.* Halstead, 136 pp.

Font-Tullot, I., 1983: *Climatología de España y Portugal* (*Climatology of Spain and Portugal*). Instituto Nacional de Meteorología, 422 pp.

Harris, J. M., 1992: An analysis of 5-day midtropospheric flow patterns for the South Pole: 1985-1989. *Tellus,* **44,** 409–421.

——, and J. D. Kahl, 1990: A descriptive atmospheric transport climatology for the Mauna Loa Observatory, using clustered trajectories. *J. Geophys. Res.,* **95,** 13 651–13 667.

Hurrell, J. W., Y. Kushnir, M. Visbeck, and G. Ottersen, Eds., 2003: An overview of the North Atlantic Oscillation. *The North Atlantic Oscillation: Climate Significance and Environmental Impact, Geophys. Monogr. Series,* No. 134, Amer. Geophys. Union, 1–35.

Jenkinson, A. F., and B. P. Collison, 1977: An initial climatology of gales over the North Sea. Met Office Synoptic Climatology Branch Tech. Memo. 62, 18 pp.

Kalkstein, L. S., G. Tan, and J. A. Skindlov, 1987: An evaluation of three clustering procedures for use in synoptic climatological classification. *J. Climate Appl. Meteor.,* **26,** 717–730.

Kanamitsu, M., 1989: Description of the NMC Global Data Assimilation and Forecast System. *Wea. Forecasting,* **4,** 335–342.

Martin, D., G. Bergametti, and B. Strauss, 1990: On the use of the synoptic vertical velocity in trajectory model: Validation by geochemical tracers. *Atmos. Environ.,* **24,** 2059–2069.

Martín-Vide, J., 1987: Característiques climatològiques de la precipitació en franja costera mediterrània de la Península Ibèrica (Climatologic features of the precipitation in the Mediterranean coast of the Iberian Peninsula). Ph.D. dissertation, Institut Cartogràfic de Catalunya, Universitat de Barcelona, 111–129.

——, 1991: *Mapas del Tiempo: Fundamentos, Interpretacíon e Imágenes de Satélite* (*Weather Maps: Fundamentals, Interpretation, and Satellite Images*). Oikos-tau, 170 pp.

——, 2001: Limitations of an objective weather-typing system for the Iberian peninsula. *Weather,* **56,** 248–250.

——, 2002: Aplicación de la clasificación sinóptica automática de Jenkinson y Collison a días de precipitación torrencial en el este de España (Application of the automatic synoptic classification of Jenkinson and Collison to days with torrential precipitations over eastern Spain). *Proc. La Información Climática Como Herramienta de Gestión Ambiental* (*Climatic Information as an Environmental Management Tool*), Zaragoza, Spain, Universidad de Zaragoza, 123–127.

Mattis, I., 2001: WP4, Compilation of trajectory data. EARLINET: A European Aerosol Research Lidar Network to Establish an Aerosol Climatology. J. Bösenberg, Ed., Max-Planck-Institute für Meteorologie Contract EVR1-CT1999-40003, 26–29. [Available online at http://lidarb.dkrz.de/earlinet/scirep1.pdf.]

Millán, M. M., R. Salvador, E. Mantilla, and G. Kallos, 1997: Photooxidant dynamics in the Western Mediterranean basin in summer: Results from European research projects. *J. Geophys. Res.,* **102,** 8811–8823.

Moody, J. L., 1986: The influence of meteorology on precipitation chemistry at selected sites in the Eastern United States. Ph.D. dissertation, University of Michigan, 176 pp.

——, and J. N. Galloway, 1988: Quantifying the relationship between atmospheric transport and the chemical composition of precipitation on Bermuda. *Tellus,* **40,** 463–479.

——, and P. J. Samson, 1989: The influence of atmospheric transport on precipitation chemistry at two sites in the midwestern United States. *Atmos. Environ.,* **23,** 2117–2132.

——, S. J. Oltmans, H. Levy II, and J. T. Merrill, 1995: Transport climatology of tropospheric ozone: Bermuda, 1988-1991. *J. Geophys. Res.,* **100,** 7179–7194.

Petisco, E., and J. M. Martín, 1995: Caracterización de la circulación general atmosférica en la Península Ibérica y Baleares (Characterization of the atmospheric general circulation in the Iberian Peninsula and Balearic Islands). Service of Analysis and Research of Climate, Instituto Nacional de Meteorología Tech. Note 5, 83 pp.

Querol, X., A. Alastuey, S. Rodriguez, F. Plana, C. R. Ruiz, N. Cots, G. Massagué, and O. Puig, 2001: PM10 and PM2.5 source apportionment in the Barcelona Metropolitan area, Catalonia, Spain. *Atmos. Environ.,* **35,** 6407–6419.

Ribalaygua, J., and R. Borén, 1995: Clasificación de patrones espaciales de precipitación diaria sobre la España Peninsular y Baleárica (Classification of daily spatial patterns of precipitation of peninsular Spain and Balearic Islands). Service of Analysis and Research of Climate, Instituto Nacional de Meteorología, Tech. Note 3, 48 pp.

Rodríguez, S., X. Querol, A. Alastuey, G. Kallos, and O. Kakaliagou, 2001: Saharan dust contributions to PM10 and TSP levels in Southern and Eastern Spain. *Atmos. Environ.,* **35,** 2433–2447.

Rolph, G. D., cited 2003: READY: Real-time Environmental Applications and Display System. NOAA Air Resources Laboratory. [Available online at http://www.arl.noaa.gov/ready.html.]

Sánchez, J., 1993: *Situaciones Atmosféricas en España* (*Atmospheric Situations in Spain*). Ministerio de Obras Públicas, Transportes y Medio Ambiente, Centro de Publicaciones, 285 pp.

Sicard, M., C. Pérez, A. Comerón, J. M. Baldasano, and F. Rocadenbosch, 2003: Determination of the mixing layer height from regular lidar measurements in the Barcelona area. *Proc. 10th Int. Symp. on Remote Sensing,* Barcelona, Spain, SPIE, Vol. 5235, 505–516.

Sinnott, R. W., 1984: Virtues of the Haversine. *Sky Telescope,* **68,** 159.

Soriano, C., J. M. Baldasano, W. T. Buttler, and K. R. Moore, 2001: Circulatory patterns of air pollutants within the Barcelona air basin in a summertime situation: Lidar and numerical approaches. *Bound.-Layer Meteor.,* **98,** 33–55.

Spellman, G., 2000: The application of an objective weather-typing system to the Iberian peninsula. *Weather,* **55,** 375–385.

Stohl, A., 1998: Computation, accuracy and applications of trajectories—A review and bibliography. *Atmos. Environ.,* **32,** 947–966.

——, and P. Seibert, 1998: Accuracy of trajectories as determined from the conservation of meteorological tracers. *Quart. J. Roy. Meteor. Soc.,* **124,** 1465–1484.

——, and T. Trickl, 1999: A textbook example of long-range transport: Simultaneous observation of ozone maxima of stratospheric and North American origin in the free troposphere over Europe. *J. Geophys. Res.,* **104,** 30 445–30 462.

Toll, I., and J. M. Baldasano, 2000: Modeling of photochemical air pollution in the Barcelona area with highly disaggregated anthropogenic and biogenic emissions. *Atmos. Environ.,* **34A,** 3069–3084.